

①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENTAMT

⑫ Offenlegungsschrift  
⑪ DE 3901675 A1

⑳ Aktenzeichen: P 39 01 675.7  
㉑ Anmeldetag: 21. 1. 89  
㉒ Offenlegungstag: 27. 7. 89

⑤① Int. Cl. 4:  
C07 H 21/04  
C 12 P 19/34  
C 12 N 15/00  
G 01 N 33/68

Behörden-  
Stempel

DE 3901675 A1

③① Unionspriorität: ③② ③③ ③①  
21.01.88 US 146508

⑦① Anmelder:  
United States Department of Energy, Washington,  
D.C., US

⑦④ Vertreter:  
Wagner, K., Dipl.-Ing.; Geyer, U., Dipl.-Phys.  
Dr.rer.nat., Pat.-Anwälte, 8000 München

⑦② Erfinder:  
Stodolsky, Marvin, Washington, D.C., US

⑤④ Reinigung polymorpher Komponenten komplexer Genome

Ein Verfahren zur Verarbeitung von in Beziehung stehen-  
den oder verwandten Subjekt- und Referenz-Makromolekü-  
len, bestehend aus einem komplementären Strang in ihren  
entsprechenden Subjekt- und Referenz-Populationen reprä-  
sentativer Fragmente und Ausführung einer Reinigung von  
einzigartigen polymorphen Subjekt-Fragmenten.

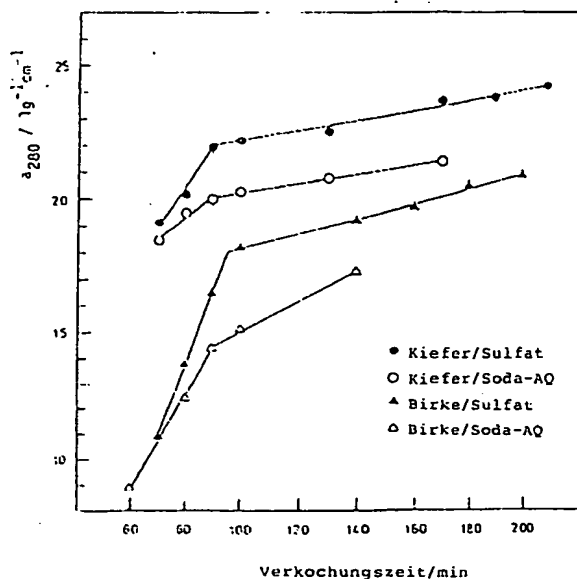


Fig. 1: Änderung der Extinktion der Ligninfraktionen

DE 3901675 A1

## Beschreibung

Die Erfindung bezieht sich allgemein auf ein Verfahren zur Charakterisierung von Makromolekülen, bestehend aus komplementären Strängen. Speziell betrifft die Erfindung ein Verfahren für subtrahierende Vergleiche von Populationen von repräsentativen Fragmenten (im folgenden PRF's genannt), die zwei in Beziehung stehende komplexe Makromoleküle repräsentieren, wie beispielsweise genomisches DNA und RNA und partielle Reinigung der polymorphen PRF-Komponenten, hinsichtlich welcher sich die zwei Makromoleküle unterscheiden.

Polymorphismen sind genetische Unterschiede zwischen zwei in Beziehung stehenden (verwandten) Genomen, die vererbbar sind und die zur Diversität oder Verschiedenheit innerhalb einer Art oder Spezies beitragen. Sie entsprechen strukturellen Subunit-(Untereinheits) Unterschieden in den DNA's (oder RNA's), welche die Genome codieren. Viele DNA-Polymorphismen sind ohne manifeste physiologische Effekte, wohingegen andere die kausalen Faktoren für ererbte Eigenschaften sind, unabhängig davon, ob die Effekte positiv, neutral oder für genetische Erkrankung die Ursache sind. Die Isolierung von Fragmenten der gesamten genomischen DNA, welche Polymorphismusplätze repräsentiert, ist eine wichtige Aufgabe der biologischen und medizinischen Forschung. Für die medizinische Genetik bilden diese Fragmentisolationen einen Schritt hinsichtlich der Kapazitätsentwicklung zur Diagnose genetischer Krankheiten. Allgemeiner ist es ein gemeinsamer Bestandteil biologischer Forschungsprogramme, Gene zu isolieren und ihre Funktionen zu kennzeichnen.

Bislang war die Detektion genetischer Unterschiede in genomischem DNA und die Isolierung von Genen durch die Kompliziertheit der Genome beschränkt, die durch konventionelle Verfahren analysiert werden konnten, ohne daß man auf schwierige arbeitsintensive vergleichende Probenahmenverfahren zurückgriff. Im folgenden werden einige dieser Verfahren und ihre Nachteile diskutiert.

Die Subtraktionshybridisation war eine der ersten Möglichkeiten zur Isolation von Genen oder ihrer entsprechenden RNA. Dieses Verfahren vertraut auf die duplex- oder doppelsträngige Struktur der DNA- und RNA/DNA-Hybride. DNA-Duplexe können denaturiert werden, d. h. getrennt werden in ihre komplementären Stränge durch die Behandlung mit Wärme oder mit Destabilisierungsagenzien, wie beispielsweise einem Formamid oder einer einen hohen pH-Wert besitzenden Lösung. Erwärmungsbedingungen können vorgesehen werden, bei denen die Stränge sich paaren und wiederum Duplexe bilden. Die Stabilität der Duplexe ist außerordentlich abhängig von der richtigen Paarung der Bestandteilsbasen über die Stränge hinweg. Die vier Bestandteilsbasen, die in DNA-Molekülen gefunden wurden, sind die folgenden: Adenin, Thymin, Guanin und Cytosin (im folgenden mit A bzw. T bzw. G bzw. C abgekürzt). Richtige Untereinheitspaarungen über die Stränge hinweg sind A mit T und G mit C. In den ersten Subtraktions-Hybridisationsexperimenten wurde virale Subjekt-DNA und Gastzellen-DNA verwendet. Die virale Komponente der gesamten extrahierten RNA aus den virusinfizierten Zellen war selektiv mit viralen, aber nicht mit Gastzellen-DNA gebunden. (Bautz und Hall, The Isolation of T4-Specific RNA on a DNA-Cellulose Column, 48 Pro. Nat. Acad. Sci. 400 [1962]). Die Hybridisierung erfolgt zwischen zwei komplementären Einzelsträngen, selbst wenn einer der Stränge stabil an einer Matrix angebracht ist.

Während der konventionellen Subtraktions-Hybridisation werden DNA's von einem Subjektgenom und einem in Beziehung stehenden (verwandten) Referenz- oder Bezugsgenom verwendet. Die Duplex DNA's von beiden werden fragmentiert und sodann denaturiert. Die fragmentierten Referenz- oder Bezugsstränge werden mit einer Matrix verbunden, wie beispielsweise Agarose, Cellulose oder Nylon. Die fragmentierten Subjektstränge werden mit einem großen Molarüberschuß der gebundenen Referenzstränge erwärmt. Während des Erwärmungsverfahrens bilden die meisten der Subjektstränge Paare mit Referenz- oder Bezugskomplementen und sie werden in Hybridduplexen von Subjekt- und Referenzsträngen eingefangen. Subjektstränge ohne Referenzkomplemente können sich nicht in einem stabilen Duplex mit den Referenz DNA's, um dadurch eingefangen zu werden, paaren. Nach dem Schritt des Erwärmens eliminiert die Entfernung der Matrix die Referenz-DNA's und die eingefangenen homologen Subjekt-DNA's. Die freien Subjekt-DNA's weisen die gesuchten einzigartigen Subjekt-DNA's und übliche DNA's auf, welche der Einfangung in Hybridduplexen entkommen sind. Die ersteren weisen eine viel größere Proportion an freien DNA's auf, als dies für die Eingangs-Subjekt-DNA-Population galt, da die Mehrheit der üblichen DNA's heraussubtrahiert wurde. Der Netto-Subtraktions-Hybridisierungsprozeß sieht somit eine partielle Reinigung oder Purifikation für die gesuchten DNA's, denen Bezugskomplemente fehlen, vor.

Das Ausmaß der Elimination unerwünschter Subjekt-DNA's während eines konventionellen Subtraktions-Hybridisations-Prozesses hängt von dem Molarverhältnis der Eingangsmaterialien ab. Die Erwärmung der Stränge in Duplexe ist eine bimolekulare Reaktion, die üblichen chemischen Massenwirkungsgesetzen genügt. Mit einem Eingangsverhältnis von einer Subjekt-DNA zu zehn Referenz-DNA's liegen die erwärmten Produkte im Verhältnis von 0,1 (Subjekt) : 2 (Hybrid) : 9 (Referenzduplexe) vor. Somit eliminiert bezüglich der Eingangs-subjekt-DNA-Population, die Elimination von matrixgebundenen DNA's 90% der Subjekt-DNA mit Referenz-Homologien.

Die konventionelle Subtraktions-Hybridisations-Technologie besitzt eine begrenzte Anwendungsfähigkeit, d. h. Polymorphismen entsprechend Weglassungen in den Genomen von einfachen Organismen, wie Viren und Bakterien. Das Verfahren ist nicht erfolgreich für Punktmutationen und Wiederanordnungs-polymorphismen. Die gesuchten Subjekt-DNA-Polymorphismen haben noch immer Homologien mit Bezugs- oder Referenz-DNA's und würden infolgedessen eingefangen und eliminiert werden während eines Subtraktions-Hybridisationsverfahrens. Darüber hinaus enthält genomisches DNA höhere Spezies zahlreiche Basenpaarsequenzen, die wiederholt werden und über die Chromosomen hinweg verteilt sind. Beispielsweise weist ungefähr 80% menschliches genomisches DNA mehrere Familien von wiederholten (reiterierten) DNA-Sequenzen auf, wobei die größten Familien Hunderte bis Tausende von Kopien besitzen. Einzelkopiegene oder Sequenzen weisen die

wünschte Komplikation. Während eines Erwärmungsvorgangs von DNA-Strängen eines komplexen Genoms machen die reiterierten Sequenzen schnellere Kontakte als die eine wesentlich niedrigere Konzentration aufweisenden Einzelkopiesequenzen. Infolgedessen bilden reiterierte Regionen stabile Duplexregionen, unabhängig von der Nicht-Homologie zwischen benachbarten Einzelkopiegenregionen. Infolgedessen bilden sich ausgedehnte "promiskose" Verknüpfungen von DNA-Form, die durch die Duplexregionen stabilisiert sind. Die Bildung von promiskosen Verknüpfungen behindert die Reinigung in der konventionellen Subtraktion-Hybridisation.

Alternative Möglichkeiten zur konventionellen Subtraktionshybridisation verwenden Restriktionsnukleasen. Eine Restriktionsnuklease ist ein Enzym, das die Kapazität besitzt, eine spezifische Target- oder Zielsequenz zu erkennen, welches mehrere Basenpaare lang ist in doppelsträngigen DNA-Molekülen und beide Stränge des DNA-Moleküls an den Stellen der Targetsequenz trennt. Die durch Digestion mit einer Restriktionsnuklease definierten DNA-Moleküle werden als Restriktionsfragmente bezeichnet. Jede gegebene genomische DNA digestiert oder "verdaut" durch eine spezielle Restriktionsnuklease, wird in ein diskretes PRF verwandelt.

Ein Restriktionsfragment-Längenpolymorphismus (im folgenden als RFLP bezeichnet) ist eine spezielle Art eines Polymorphismus, manifestiert als ein Längenunterschied einiger genetisch verwandter Fragmente der zwei verglichenen PRF's. Die zugrunde liegenden genetischen Manifestationen können so fein wie eine einzige Basenpaaränderung sein, welche eine Schneidstelle schafft oder eliminiert oder so grob oder deutlich wie eine genetische Weglassung, welche die Länge der DNA zwischen Spalt-, Trenn- oder Schneidplätzen ändert. Um einen RFLP zu detektieren, ist ein analytisches Verfahren zur Fraktionierung doppelsträngiger DNA-Moleküle auf der Basis der Größe erforderlich. Das gebräuchlichste Verfahren zur Erreichung einer solchen Fraktionierung ist die Agarosegel-Elektrophorese. Bei diesem Verfahren wandern die DNA-Moleküle durch das Gel, welches als ein Sieb wirkt, das die Bewegung der größten Moleküle im größten Ausmaß verzögert und die Bewegung der kleinsten Moleküle im kleinsten Ausmaß beeinflusst. Ein Vergleich der durch Gel-Elektrophorese fraktionierten PRF's zeigt, daß die Fragmente einzigartig für jedes Genom unter denen gemeinsam für die Subjekt- und Referenz-PRF's verglichen werden. Die einzigartigen Fragmente repräsentieren den RFLP. Die fraktionierten PRF's können auch denaturiert und innerhalb der Grenzen des Fraktionierungsgels angelassen oder erwärmt werden. Solche "in situ"-Erwärmungen wurden zuvor verwendet, und zwar in einer Strategie zum selektiven Detektieren von reiterierten PRF-Gliedern. (Roninson, Detection and mapping of homologous, repeated and amplified DNA sequences by DNA renaturation in agarose gels, *m 11 Nucleic Acids Res.* 5413.31 [1983]). Fraktionierungen, welche die verglichenen DNA's durch die Stabilität der Basenpaarung unterscheiden, wurden ebenfalls verwendet, (Fischer and Lerman, Length-Independent Separation of DNA Restriction Fragments in Two-Dimensional Gel Electrophoresis, 16 *Cell* 191-200 [Jan. 1979]). Sie können einen gewissen Polymorphismus zwischen DNA's der gleichen Länge zeigen.

Solange ein Fraktionierungsverfahren die Bestandteile jedes PRF auflöst, werden die Differenzen zwischen PRF's leicht detektierbar. Beispielsweise kann die gewünschte Auflösung mit eindimensionalen Fraktionierungen für virale PRF's erreicht werden oder mit zwei-dimensionalen Fraktionierungen, ansprechend auf die Fragmentlänge und die thermische Stabilität für bakterielle PRF's. Für höhere Organismen jedoch wird selbst, wenn beste Fraktionierungsverfahren verwendet werden, die gewünschte Auflösung polymorphischer PRF-Bestandteile nicht erreicht. Bei solchen höheren Organismen tritt die Trennung jedes einzelnen Glieds von der Majorität der PRF-Mitgliedschaft auf, aber es gibt so viele Glieder, daß ein Kontinuum aus überlappenden Fragmentbändern vorhanden ist, welche die Auflösung und Detektion von Gliedern innerhalb des Kontinuums verhindern.

Wenn ein Kontinuum von Fragmentbändern vorhanden ist, wurden Probier- oder Probennahmenverfahren verwendet, um Positionen spezieller Gene darzustellen. Einer geklonten Form des Gens, welches gesucht wird, wird eine radioaktive oder biochemische Markierung (Etikett oder "label") gegeben, das später verwendet werden kann, um die Position zu enthüllen. Es dient als eine Probe oder Sonde zur Lokalisierung seiner Homologe. Die fraktionierten Subjekt-DNA's sind denaturiert in Bestandteilsstränge und werden sodann transferiert und stabil gebunden an eine Membran, beispielsweise abgezogen oder abkopiert auf eine stabile Membran. Eine einzelsträngige Probe und die kopierten Subjekt-DNA's werden sodann erwärmt. Die Probe bindet in stabiler Art und Weise durch Basispaarung nur an der Position ihrer genetischen Homologe und die Positionen der Homologenfragmente auf der Kopie, wodurch die Detektion stattfindet. Mit den meisten Einzelgenproben zeigen die PRF's keine Differenzen für die selektiv gezeigten Fragmente. Nichtsdestoweniger können mühevollen Vergleichsprobenuntersuchungen von in Beziehung stehenden oder verwandten PRF's sequentiell ausgeführt werden und mit einer Population von Proben, die groß genug ist, können schließlich Polymorphismen detektiert werden, die für genetische diagnostische Zwecke brauchbar sind. (Gusella et al., A Polymorphic DNA Marker Genetically Linked to Huntington's Disease, 306 *Nature* 234 [1983]).

Ein weiteres Verfahren wurde verwendet zur selektiven Darstellung einer Sub-Population von Polymorphismen von viralem genomischem DNA. Bei diesem Verfahren werden PRF's der genomischen DNA von zwei Genomen, die verglichen werden sollen, hergestellt. Sie werden mit gleichen Mengen gepoolt oder zusammengebracht und hybridisiert. Die Hybridisationsprodukte werden sodann mit Nuklease S1 behandelt, welches an den Verformungen in den DNA-Duplexen trennt. (Shenk et al., Biochemical Method for Mapping Mutational Alterations in DNA with S1 Nuclease: The Location of Deletions and Temperature-Sensitive Mutations in Simian Virus 40, 72 *Proc., Nat. Acad. Sci.* 3 : 989-993 [1975]). Einige Hybridduplexe, die aus polymorphen DNA-Strängen bestehen, besitzen ein hinreichendes Ausmaß an Verformung und werden infolgedessen an diesen Stellen oder Plätzen getrennt. So erzeugte sekundäre Fragmente werden durch eine Fraktionierung detektiert, während welcher die S1-Trennungsfragmente schneller wandern als die intakten Vorgängerfragmente. Für dieses Verformungstrennverfahren ist es wichtig, daß eine Steuerung, bestehend aus einer Hybridisation jedes PRF gegen sich selbst für eine komparative Analyse der Produkte ausgeführt wird. Solche Steuerungen ergeben sekundäre S1-Fragmente, die wegen der partiellen Homologien und reiterierten Sequenzen innerhalb

der genomischen DNA auftreten. Die Fragmente, welche diese codieren, bilden verformte Duplexe und partielle Duplexkomplexe während der Hybridisierungen. Somit entstehen sekundäre Fragmente während der SI-Nukleasedigestion oder Verdauung. Diese sekundären Fragmente müssen identifiziert werden, um Polymorphismen zwischen Genomen von internen Homologien mit einem Genom zu unterscheiden. Diese Verformungstrennungstechnologien wurden auch mit bakteriellem genomischem DNA verwendet. (Yee and Inouye, Two-dimensional SI nuclease heteroduplex mapping: Detection of rearrangements in bacterial genomes, 81 Proc. Nat. Acad. Sci. 2723—2727 [1984]). Interne Homologien sind ein kleiner Anteil des gesamten genomischen DNA in bakteriellen Genomen und sie sind von der Steuerung identifizierbar. Im Gegensatz dazu sind die internen Homologien in dem genomischen DNA höherer Organismen extensiv. Wenn infolgedessen dieses Verfahren mit PRF's höherer Organismen verwendet wird, werden die gesuchten Polymorphismen durch die große Reichhaltigkeit an sekundären Fragmenten überdeckt, welche sich als eine Folge der extensiven internen Homologien ergibt.

Andere Polymorphismus-Identifikationsverfahren wurden mit einem begrenzten Sicherheitsbereich verwendet. Es handelt sich dabei um Verfahren, welche das vorausgehende Clonen des Genomfragmentes erforderlich machen, dessen Polymorphismus darauffolgend gesucht wird. Das verfeinerte Verfahren dieser Methoden ist die vergleichende Nucleotid-Sequenzierung, durch welche spezielle Untereinheit (subunit)-Unterschiede des Polymorphismus identifiziert werden.

Kurze Zusammenfassung der Erfindung. Ein Ziel der Erfindung besteht darin, ein neues Verfahren zur Detektion von mindestens einem Unterschied zwischen zwei in Beziehung stehenden Makromolekülen festzustellen, und zwar bestehend aus komplementären Strängen (im folgenden Makromoleküle genannt), wie beispielsweise Duplex DNA und Duplex RNA. RNA = RNS; DNA = DNS.

Ein weiteres Ziel der Erfindung besteht darin, ein neues Verfahren anzugeben, um die gleichzeitige partielle Reinigung vieler einzigartiger Glieder der PRF von Subjekt-Makromolekülen zu erhalten, welchen Komplemente in den Fragmenten der Referenz-Makromoleküle fehlen.

Ein drittes Ziel der Erfindung besteht darin, ein neues Verfahren vorzusehen zur Identifizierung von extrinsischen Additionen Zu- und Wiederanordnungen innerhalb eines subjektgenomischen DNA's, verglichen mit einem entsprechenden referenzgenomischen DNA.

Ein Merkmal der vorliegenden Erfindung ist ein Verfahren, welches die partielle Reinigung ermöglicht und die darauffolgende Detektion einer Klasse von einzigartigen Gliedern des PRF des genomischen Subjekt-DNA von Gliedern, die den Subjekt- und Referenz-PRF's gemeinsam sind, und zwar selbst dann, wenn die PRF's so komplex sind, daß die Fraktionierung selbst nicht die Bestandteiglieder innerhalb jedes PRF auflöst. Ein einziges PRF-Glied ist der Satz von genetisch identischen Fragmenten, entsprechend identischen Segmenten des mehrfach identischen Substratgenome, erzeugt durch die platzspezifischen Trennungen der Substratgenome. Die Klasse einzigartiger Glieder der verglichenen PRF's wird durch das spezielle gewählte Fraktionierungsverfahren bestimmt. Derartige Fraktionierungsverfahren würden beispielsweise folgendes umfassen: Trennung durch Fragmentlänge, Durchschnittsuntereinheit-Zusammensetzung, Einleitung der Doppelhelix in Einzelstrangübergang, Beendigung der Doppelhelix auf einen einzigen Strangübergang und Kapazität zur Bindung jeder Verschiedenheit von Agenzien. Die gewählte Primärfraktionierung erreicht eine Trennung jedes speziellen PRF-Glieds vom größten Teil der begleitenden Eingangsgröße in die Fraktionierung. PRF-Glieder von entsprechenden genetischen Orten der subjekt- und referenzgenomischen DNA, die nicht kofraktionieren werden sich in der Menge von polymorphen Gliedern, gereinigt durch das erfindungsgemäße Verfahren befinden. Wenn beispielsweise Glieder, abgeleitet aus entsprechenden Orten für die subjekt- und referenzgenomische DNA sich in der Länge unterscheiden (d. h. RFLP) sind, so ergibt das erfindungsgemäße Verfahren Subjektglieder, welche die RFLP-Orte repräsentieren.

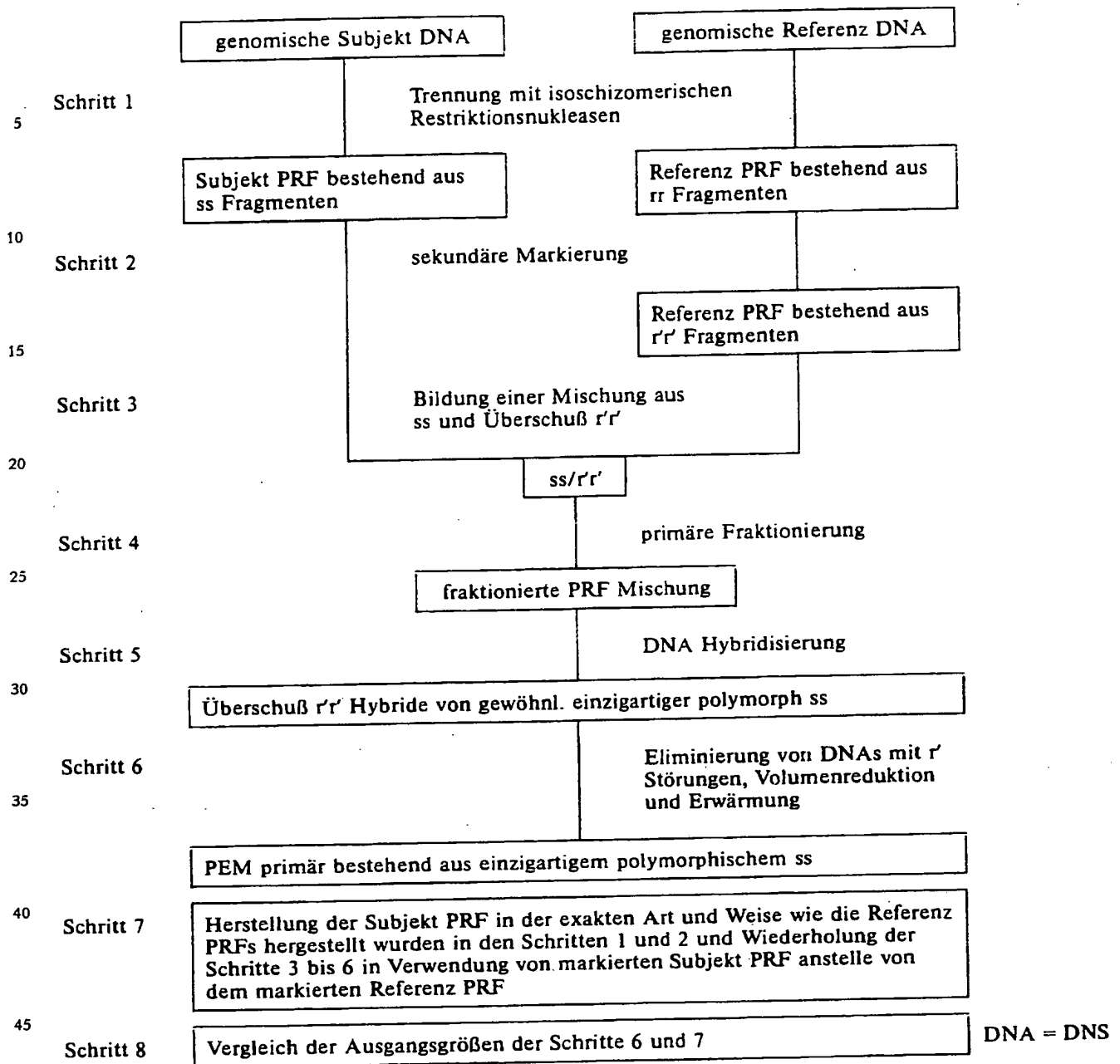
Die Erfindung ist auch in der Lage, die Probleme zu überwinden, die durch einen Überfluß an wiederholten Sequenzen auftreten, die charakteristisch für höhere eukaryotische Organismen (Organismen, deren Zellen Kerne enthalten) sind. Dies wird erreicht durch Ausführung der obigen Fraktionierung vor der subtraktiven Hybridisierung. Die Fraktionierung verteilt die PRF-Glieder einschließlich derjenigen mit wiederholten Folgen in zahlreiche unterschiedliche Fraktionen. Mit der Kompliziertheit jeder Fraktion, die wesentlich kleiner ist als die der gesamten Eingangs-PRF's, wird das Potential für promiskue Komplexbildung innerhalb jeder Fraktion entsprechend vermindert während der darauffolgenden subtraktiven Hybridisation. Die Erfindung ist in der Lage, die Probleme des zufälligen DNA Brechens während des obigen Verfahrens zu überwinden, und zwar durch den Einschluß von Schritten, abhängig von dem Vorhandensein von ursprünglichen Paaren von Fragmenten, erzeugt während der Reduktion der genomischen DNA's in ihre PRF's.

In der bevorzugten Form der Erfindung richtet sich das Verfahren allgemein auf (1) die Umwandlung des Subjekt- und Referenz-Genome in ihre entsprechenden PRF's, (2) das Vorsehen der Subjekt- und Referenz-PRF's mit unterschiedlichen biochemischen und/oder isotopischen Markierungen; (3) die Bildung einer Mischung der Subjekt-PRF und der Referenz- oder Bezugs-PRF; (4) die Fraktionierung der Mischung; (5) die Denaturierung der Fragmente innerhalb jeder Fraktion in einzelne DNA-Stränge und die Erwärmung der Stränge zur Wiederbildung der Duplexe, die folgendes aufweisen (a) Hybidduplexe von Strängen gemeinsamer Sequenz, sowohl für die Subjekt- wie auch die Referenz-PRF's, (b) Restsubjektfragmente, die einzigartig für die Subjekt-PRF sind, und (c) Überschußreferenz-Fragmente; (6) Verwendung distinktiver biochemischer und/oder isotopischer Markierungen für die Reinigung der Subjektfragmente, die nicht in Duplexen eingefangen wurden, von ihren Referenz-Homologen, wodurch die gewünschte partielle Reinigung der Fragmente einzigartig für die Subjekt-PRF erreicht wird; (7) Ausführung einer Steuerreinigung an Subjekt-PRF alleine und (8) Vergleich des Produkts des Schrittes (7) mit dem Produkt des Schrittes (6) zur Identifizierung von noch immer vorhandenen Subjekt-sequenzähnlichen Subjekt-Fragmenten in dem teilweise gereinigten Produkt der Schritte (1) bis (6). Andere

20. Verfahren zur Verarbeitung von in Beziehung stehender genomischer Subjekt- und Referenz-DNA in deren entsprechende Subjekt- und Referenz-PRF's und Bewirkung der partiellen Reinigung von einzigartigen polymorphen Subjekt-Restriktionsfragmenten, wobei die folgenden Schritte vorgesehen sind:

- (a) Umwandlung der genomischen Subjekt- und Referenz-DNA in entsprechende Subjekt- und Referenz-PRF's,
- (b) Vorsehen unterscheidender biochemischer und/oder isotopischer Markierungen für die Referenz-PRF,
- (c) Bildung einer Mischung des markierten Referenz-PRF und des Subjekt-PRF,
- (d) Fraktionierung von Gliedern der PRF's in der Mischung,
- (e) Durchführung einer DNA-Hybridisierung der fraktionierten Mischung,
- (f) Verwendung der unterscheidenden biochemischen und/oder isotopischen Markierungen für die Trennung der Referenz-PRF's und der DNA-Hybride der Subjekt- und Referenz-PRF's von den polymorphen Gliedern des Subjekt-PRF und Erreichung der partiellen Reinigung der polymorphen Glieder des Subjekt-PRF und Anlassen jedweder verbleibender einzelner Subjekt-Stränge in der partiellen Reinigung in Duplexe,
- (g) Herstellung von Subjekt-PRF in der exakten Art und Weise wie das Referenz-PRF hergestellt wurde in den Schritten (a) und (b) und Wiederholung der Schritte (c) bis (f) unter Verwendung des markierten Subjekt PRF anstelle des markierten Bezugs-PRF,
- (h) Vergleich der Ausgangsgrößen der Schritte (f) und (g) zur Durchführung des Schritts der Erkennung des Polymorphismus innerhalb der partiell gereinigten polymorphen Subjekt-Fragmente,
- (i) weitere Reinigung der polymorphen Glieder des Subjekt-PRF zur analytischen Verwendung unter Benutzung von einer oder zwei Dimensionsfraktionierungen zu Anzeigezwecken, Bildung von Rekombinations-DNA und genetischer Verstärkung unter Verwendung rekombinierender DNA-Cloning, und/oder Wiederholung der Schritte (a) bis (h).

2. Verfahren nach Anspruch 1, wobei die unterscheidenden Markierungen biochemische und/oder isotopische Markierungen aufweisen.
3. Verfahren nach Anspruch 1 oder 2, wobei die Makromoleküle zwei Duplex genomische DNA's aufweisen.
4. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, insbesondere nach Anspruch 1, dadurch gekennzeichnet, daß die Makromoleküle zwei doppelsträngige RNA aufweisen. 5
5. Verfahren nach einem oder mehreren der vorhergehenden Ansprüche, insbesondere nach Anspruch 1, dadurch gekennzeichnet, daß die Makromoleküle aus der Gruppe ausgewählt sind, die im wesentlichen aus genomischem DNA und doppelsträngigem RNA besteht.
6. Verfahren zur Verarbeitung von in Beziehung stehendem Subjekt- und Referenz-genomischem DNA in ihrer entsprechenden Subjekt- und Referenz-PRF's und Bewirkung einer partiellen Reinigung von einzigartigen polymorphen Subjekt-Restriktionsfragmenten, wobei die folgenden Schritte vorgesehen sind: 10
  - (a) Umwandlung des genomischen Subjekt- und Referenz-DNA in entsprechende Subjekt- und Referenz-PRF's,
  - (b) Vorsehen unterscheidender biochemischer und/oder isotopischer Markierungen für das Referenz-PRF, 15
  - (c) Bildung einer Mischung des markierten Referenz-PRF und des Subjekt-PRF,
  - (d) Fraktionierung der Glieder der PRF's in der Mischung,
  - (e) Durchführung einer DNA-Hybridisierung der fraktionierten Mischung,
  - (f) Verwendung der unterscheidenden biochemischen und/oder isotopischen Markierungen zur Trennung der Referenz-Duplexe und der DNA-Hybride der Subjekt- und Referenz-PRF's aus den polymorphen Gliedern des Subjekt-PRF und Erreichung der partiellen Purifikation der polymorphen Glieder der Subjekt-PRF's und Erwärmen jedweder verbleibender einzelner Subjekt-Stränge in der partiellen Reinigung oder Purifikation in Duplexe, 20
  - (g) Herstellung von Subjekt-PRF in der genauen Art und Weise wie das Bezugs-PRF hergestellt wurde in den Schritten (a) und (b) und Wiederholung der Schritte (c) bis (f) unter Verwendung des markierten Subjekt-PRF anstelle des markierten Referenz-PRF, und 25
  - (h) Vergleichen der Ausgangsgröße der Schritte (f) und (g) zur Durchführung des Schrittes der Erkennung des Polymorphismus innerhalb der partiell gereinigten polymorphen Subjekt-Fragmente.
7. Verfahren nach Anspruch 6, wobei der Schritt der Umwandlung der Subjekt- und Referenz-Genom-DNA ein entsprechendes erwähntes PRF das Schneiden der erwähnten genomischen DNA mit irgendeiner konventionellen Restriktionsnuklease umfaßt. 30
8. Verfahren nach Anspruch 6, wobei der Schritt der Umwandlung des genomischen Subjekt- und Referenz-DNA ein entsprechendes erwähntes PRF das Schneiden des genomischen DNA umfaßt, um definierte Fragmente im Gegensatz zu zufällig gebrochenen Strängen vorzusehen.
9. Verfahren nach Anspruch 6, wobei die genomische Subjekt-DNA in in das erwähnte Subjekt-PRF umgewandelt wird und die genomische Referenz-DNA in das erwähnte Referenz-PRF umgewandelt wird unter Verwendung eines isoschizomeren Paares von Nukleasen. 35
10. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß die genomische Subjekt-DNA in das erwähnte Subjekt-PRF umgewandelt wird unter Verwendung von Asp718 und die erwähnte genomische Referenz-DNA für den das Referenz-PRF unter Verwendung von KpnI umgewandelt. 40
11. Verfahren nach Anspruch 6, wobei die genomische Subjekt-DNA in das erwähnte Subjekt-PRF unter Verwendung von KpnI umgewandelt wird und die genomische Referenz-DNA wird unter Verwendung von Asp718 in das erwähnte Referenz-PRF umgewandelt.
12. Verfahren nach Anspruch 6, wobei die erwähnten distinktiven oder unterscheidenden biochemischen und/oder isotopischen Markierungen die selektive Entfernung des Referenz-PRF aus der Mischung von Subjekt- und Referenz-PRF gestatten, ohne unterschiedliche Mobilitäten auf die genetisch identischen Glieder aus Subjekt- und Referenz-PRF's während des Schrittes der Fraktionierung der PRF-Glieder zu übertragen. 45
13. Verfahren nach Anspruch 6, wobei der Schritt des Vorsehens unterscheidender Markierungen den Schritt der Verwendung von photodynamischer Biotinylation umfaßt. 50
14. Verfahren nach Anspruch 6, wobei der Schritt die Bildung der Mischung der markierten Referenz-PRF und der Subjekt-PRF die Verwendung von Mengen der Mischung umfaßt, die kompatibel ist mit dem Fraktionierungsschritt und dem Schritt jedweder Verstärkung der einzigartigen polymorphen Glieder des Subjekt-PRF.
15. Verfahren nach Anspruch 6, wobei die Mischung der markierten Referenz-PRF und der Subjekt-PRF mindestens ein 1 : 1-Verhältnis um das markierte Referenz-PRF zu dem Subjekt-PRF aufweist. 55
16. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß der Schritt der Fraktionierung die Verwendung eines Fraktionierungsverfahrens umfaßt, welches die Trennung eines speziellen Glieds der PRF's von der Majorität der anderen Bestandteile der PRF's umfaßt.
17. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß der Schritt der Fraktionierung der PRF-Glieder die gemeinsame Wanderung (co-migration) der genetisch identischen Glieder der Subjekt- und Referenz-PRF's ermöglicht. 60
18. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß der Schritt der Durchführung der DNA-Hybridisierung innerhalb jeder Fraktion gebildet während des Fraktionierungsschrittes ausgeführt wird.
19. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß der Schritt der Durchführung der DNA-Hybridisierung die Denaturierung und sodann das Erwärmen der Glieder der Subjekt- und Referenz-PRF's umfaßt, um DNA-Hybridduplexe der Subjekt- und Referenz-DNA-Stränge zu erzeugen, reine Subjekt-DNA-Duplexe, reine Referenz-DNA-Duplexe und restliche einzelne Subjekt- und Referenz-DNA-Stränge. 65



# Patentansprüche

1. Verfahren zur Verarbeitung von in Beziehung stehender Subjekt- und Referenz-Makromoleküle bestehend aus komplementären Strängen in deren entsprechende Subjekt- und Referenz-Populationen repräsentativer Fragmente (PRF's) und Bewirkung der partiellen Reinigung von einzigartigen polymorphen Subjekt-Fragmenten, **gekennzeichnet durch**

- (a) Umwandlung der Subjekt- und Referenz-Makromoleküle in entsprechende Subjekt- und Referenz-PRF's,
- (b) Vorsehen von unterscheidenden Markierungen an jedem der Referenz-PRF,
- (c) Bildung einer Mischung der markierten Referenz-PRF und der Subjekt-PRF,
- (d) Fraktionierung von Gliedern der PRF's in der Mischung,
- (e) Durchführung einer Hybridisierung der fraktionierten Mischung,
- (f) Verwendung der unterscheidenden Markierungen für die Trennung der erwähnten Referenz-Duplexe und Hybride der Subjekt- und Referenz-Fragmente von den Subjekt-Fragmenten und Erreichung der partiellen Reinigung der polymorphen Subjekt-Fragmente und
- (g) Herstellung der Subjekt-PRF in der genauen Art und Weise, wie die Referenz-PRF hergestellt wurde, in den Schritten (a) und (b) und Wiederholung der Schritte (c) bis (f) unter Verwendung der markierten Subjekt-PRF anstelle der markierten Referenz-PRF,
- (h) Vergleichen der Ausgangsgröße der Schritte (f) und (g) zur Durchführung des Schrittes der Erken-

wobei das konvektive Mischen mit benachbarten Fraktionen verhindert wird.

Folgend auf die DNA-Hybridisierung werden die unterscheidenden biochemischen und/oder isotopischen Markierungen, die den Referenz-PRF's während Schritt zwei hinzugefügt wurden, in Schritt sechs verwendet, welches der sekundäre Fraktionierungsschritt ist. Diese Markierungen ermöglichen eine Trennung der Referenz-DNA's und DNA-Hybride von Subjekt- und Referenz-DNA-Strängen von Subjekt-DNA's. Dieser Schritt eliminiert den Überschuß an Eingangsreferenz-DNA zusammen mit den Hybriden. Das Ergebnis dieser sekundären Fraktionierung ist eine partielle Reinigung der einzigartigen polymorphischen Glieder der erwähnten Subjekt-PRF. Im bevorzugten Ausführungsbeispiel werden die biotinylierten Referenz-DNA's zusammen mit ihren Hybriden mit isogenen Subjekt-Strängen durch Chromatographie über einem Harz mit angebrachtem Avidin oder Streptavidin entfernt, und zwar vertrauend auf die enge Bindung der biotinmarkierten Referenz-DNA an einer Matrix mit gebundenem Avidin. Der Anteil der gesuchten polymorphischen Subjekt-PRF-Glieder unter dem freien Subjekt-DNA wird wesentlich durch die Entfernung der gemeinsamen PRF-Komponenten erhöht. Diese Population freier Subjekt-DNA's wird bezeichnet als Polymorphismus-angereicherte Glieder (polymorphism enriched members = PEM).

Durch die obige Eliminierung der unerwünschten Referenz- und Hybrid-DNA's wird eine beträchtliche Reduktion des Volumens der verbleibenden PEM erreicht. Diese Volumenverminderung ist außerordentlich erwünscht, weil teure enzymalogische Reagenzien darauffolgend erforderlich sind, in konzentrationsabhängigen Reaktionen für die weiteren Reinigungsschritte und Rekombinations-(recombinant) DNA-Verfahren. Durch den entsprechenden Anstieg der Subjekt-DNA-Konzentration kann eine weitere Erwärmung der verbleibenden Einzelsubjekt-Stränge in Duplexe ausgeführt werden.

Es wird erwartet, daß ein Teil der gemeinsamen Subjekt-PRF-Glieder zusammen mit DNA-Verarbeitungsabfall einschließlich intakter und gebrochener Einzelstränge, gebrochener Duplexe und möglicherweise einiger promiskoser Verwirrungen den Schritt sechs des Verfahrens überstehen. Diese Verunreinigungen bilden zusammen einen reduzierten Hintergrund gemeinsamer Fragmente, innerhalb von dem die gesuchten einzigartigen Glieder erkannt werden müssen. Die Identifizierung dieses verminderten Hintergrundes wird in den Schritten sieben und acht des Verfahrens vorgesehen. Der Schritt sieben des Verfahrens umfaßt die Herstellung eines Teils des Subjekt-PRF in der gleichen Art und Weise wie das Referenz-PRF in den Schritten 1 und 2 des Verfahrens hergestellt wurde. Die auf diese Weise hergestellten Subjekt-Fragmente werden als "s's'" bezeichnet. Sodann wird eine ss/s's'-Mischung entsprechend der ss/r'r'-Mischung, gebildet in Schritt drei und verwendet in den Schritten vier bis sechs gebildet. Die s's'-Komponente der Mischung besitzt die gleichen Ausdrücke wie r'r' der ss/r'r'-Mischung in Schritt drei des Prozesses. Sodann wird diese ss/s's'-Mischung durch die gleichen Schritte vier bis sechs verarbeitet, wie dies für die ss/r'r'-Mischung geschah. Da die ss/s's'-Mischung keine einzigartigen Komponenten besitzt, ist die Ausgangsgröße dieser Wiederholung der Schritte vier bis sechs identisch zu dem gemeinsamen Hintergrund der die PEM verunreinigt. Diese Ausgangsgröße wird als Steuer- oder Kontrollfragment-Mitgliedschaft bezeichnet (im folgenden als CFM = control fragment membership). Im Schritt acht des Verfahrens werden PEM und CFM verglichen. Die Differenz zwischen dem PEM und dem CFM sind die gesuchten einzigartigen polymorphischen Glieder der Subjekt-PRF.

Das Ziel der weiteren Verarbeitung besteht im Erkennen und/oder Reinigen der einzigartigen polymorphen Glieder des PEM von dem gemeinsamen Hintergrund, repräsentiert durch die CFM. Die im Schritt eins des Verfahrens erzeugten Endmarkierungen werden verwendet zur weiteren Charakterisierung und Verarbeitung der polymorphischen Subjekt-PRF-Komponenten. Diese Kennzeichnungs- und Verarbeitungsschritte können eine komparative PEM- und CFM-Anzeige umfassen, die Bildung von rekombinierenden DNA's und die Rückführung durch die Schritte eins bis sieben mit PEM oder verstärktem PEM dienend als die Eingangsgröße.

Während der Rekombinations-DNA-Bildung und Verstärkung kann ein Vektor DNA für Rekombinations-DNA-Bildung verwendet werden mit Ausdrücken, die komplementär zu den Asp718 erzeugten Fragmentenden der ss-Fragmente sind. Solche Vektor-DNA's können nicht "Basen-paaren" mit den Hybridisierungsverunreinigungen noch dem DNA-Verarbeitungsabfall, können aber "Basen-paaren" mit und sind angebracht durch enzymatische Reaktionen an den ss-Fragmenten zur Bildung linearer Moleküle, verwendet als Ausgangsstoffe für lebensfähige Rekombinations-DNA's. Infolgedessen gestattet die Verwendung der isoschizomerischen Fragmenteendifferenzen, erzeugt im Schritt eins die Auswahl nach und die Reinigung der gesuchten polymorphischen ss-Fragmente.

Bei einer anderen Ausbildungsform der Erfindung kann das Verfahren benutzt werden, um in gleicher Weise doppel-strängige "Boten" RNA und andere Makromoleküle, bestehend aus komplementären Strängen zu verarbeiten.

Die Erfindung hat den Vorteil der Identifizierung und partiellen Reinigung vieler polymorphischer Glieder eines PRF's zu einer Zeit, selbst wenn das PRF so komplex ist, daß die konventionelle Fraktionierung selbst nicht die Bestandteiglieder innerhalb jedes PRFG auflöst. Es kann daher dazu dienen, die Genomänderungen zu detektieren, welche zu Viruserkrankheiten, Krebs und genetisch geerbten Krankheiten beitragen. Das Verfahren kann ebenfalls dienen zur Identifizierung von Genomänderungen mit Eigenschaften, die bei landwirtschaftlichen und biotechnologischen Bemühungen vorteilhaft sind.

Zusammenfassend sieht die Erfindung folgendes vor:

Ein Verfahren zur Verarbeitung von in Beziehung stehenden oder verwandten Subjekt- und Referenz-Makromolekülen, bestehend aus einem komplementären Strang in ihren entsprechenden Subjekt- und Referenz-Populationen repräsentativer Fragmente und Ausführung einer Reinigung von einzigartigen polymorphischen Subjekt-Fragmenten.



Referenz-DNA-Moleküle und Duplex-Hybride der Subjekt- und biotinylierten Referenz-DNA-Stränge, während die nicht biotinylierten Subjekt-DNA's frei während eines sekundären Fraktionierungsprozesses in Schritt sechs hindurchlaufen.

In dem dritten Schritt des Verfahrens gemäß Fig. 1 wird eine Mischung aus dem r'r'-Referenz-PRF und dem ss-Subjekt-PRF gebildet und als "ss/r'r'" bezeichnet. In diesem bevorzugten Ausführungsbeispiel sollte diese Mischung in Mengen vorliegen, die kompatibel sind mit dem Fraktionierungsverfahren, verwendet im Schritt vier und jeder späteren Verstärkung der polymorphischen Subjekt-DNA-Fragmente durch Cloning nach dem siebten Verfahrensschritt. Ein hohes Referenz-PRF zu Subjekt-PRF-Verhältnis erhöht das spätere Einfangen von Subjekt-Restriktionsfragment-Komponenten mit isogenen Referenz-Restriktionsfragment-Komponenten in Duplex DNA-Hybriden, bestehend aus Subjekt- und Referenz-DNA-Strängen. Der Ausdruck "isogen" bedeutet die Codierung der gleichen Sequenz oder Folge von genetischem Material. Isogene Fragmente können sich durch Punktmutationen (beispielsweise ein Basenpaar) unterscheiden, besitzen aber keine substantiellen Differenzen in Gehalt oder Ordnung. Dieser Schritt kann unter Verwendung eines 1 : 1-Verhältnisses von biotinyliertem Referenz-PRF zu Subjekt-PRF ausgeführt werden. Wie jedoch oben bemerkt wurde, verbessert sich der Wirkungsgrad, wenn das Verhältnis ansteigt. Beispielsweise wurde dieser Teil des Protokolls ausgeführt unter Verwendung eines 10 : 1-Verhältnisses von biotinyliertem Referenz-PRF zu Subjekt-PRF. Im bevorzugten Ausführungsbeispiel würde ein noch größerer Wirkungsgrad erwartet werden bei Verhältnissen von 100 : 1 und 1000 : 1.

Im vierten Schritt des Verfahrens wird die im dritten Schritt erzeugte Mischung fraktioniert unter Verwendung irgendeines konventionellen Fraktionierungsverfahrens, welches folgendes vorsieht (1) es trennt jedes spezielle Glied eines PRF von der großen Majorität der anderen Bestandteile dieses PRF's und (2) die gemeinsame Wanderung (co-migration) von Gliedern gestattet, und zwar genetisch identisch hinsichtlich Subjekt- und Referenz-PRF's, unabhängig von ihren unterschiedlichen Markierungen. Obwohl es wichtig ist, daß jedes Eingangsglied vom größten Teil der begleitenden Eingangsgröße während des Fraktionierungsverfahrens getrennt wird, ist es nicht notwendig, daß physikalisch überlappende Zonen von PRF-Gliedern während der Fraktionierung aufgelöst werden. Diese Fraktionierung definiert zwei Klassen von Subjekt-Gliedern. Die üblichen Subjekt-Glieder sind diejenigen, die gemeinsam wandernde und isogene Referenzpartner besitzen. Die einzigartigen Subjektglieder sind diejenigen polymorphischen Subjekt-DNA-Fragmente, denen isogene und gemeinsam wandernde Referenzpartner fehlen. Es sind diese einzigartigen polymorphischen Fragmente, die durch das Netz gereinigt werden, d. h. den gesamten Prozeß.

Die substantive Trennung jedes Glieds vom größten Teil der begleitenden Eingangsgröße erreicht ein zweites wichtiges Ziel. Die Trennungen vermindern die Wahrscheinlichkeit, daß irgendein Glied koresidente Glieder mit akzidentalen (zufälligen) Homologen aufweist, beispielsweise reiterierte Sequenzen. Auf diese Weise wird die Bildung von promiscösen Störungen in der darauffolgenden DNA-Hybridisierung stark vermindert. Jedes Glied koresidiert mit viel weniger Gliedern als in der anfänglichen PRF-Eingangsgröße.

Das bevorzugte Verfahren zur Erreichung der Fraktionierung der Mischung ist die Größenfraktionierung durch Elektrophorese durch ein Agarosegel. Bei diesem Verfahren wandern DNA-Moleküle durch das Gel, als ob es ein Sieb wäre, das die Bewegung der größten Moleküle im größten Ausmaß verzögert und die Bewegung der kleinsten Moleküle im geringsten Ausmaß verzögert. Daher ist die Beweglichkeit bei Elektrophorese in dem Agarosegel um so größer, je kleiner das Restriktionsfragment ist.

Glieder, die Längenpolymorphismus in dem Subjekt-PRF zeigen, fraktionieren nicht gemeinsam mit den partiellen Homologen des Referenz-PRF, wohingegen isogene Subjekt- und Referenz-Glieder kofraktionieren. Das Fraktionierungsverfahren kann entweder mit den DNA's noch eingefangen innerhalb des Gels enden oder durch Sammlung von Fraktionen von Abläufen vom Verfahren. Die Beibehaltung der DNA's innerhalb des Gels bewahrt am besten die Gliedertrennungen, die durch das Fraktionierungsverfahren erreicht wurden. Der Ausdruck Fraktion wird auf die Zonen des Gels angewandt mit ihrer eingefangenen Mitgliedschaft und auch auf die gesammelten Fraktionen von aus der Fraktioniervorrichtung austretenden Ablauf.

Der fünfte Schritt des in Fig. 1 gezeigten Verfahrens sieht die Durchführung einer DNA-Hybridisierung vor, und zwar innerhalb jeder Fraktion, gebildet während des Schrittes vier. Während der DNA-Hybridisierung werden die Restriktionsfragmente zuerst denaturiert, d. h. die Duplex-Restriktionsfragmente werden in ihre komplementären einzelnen DNA-Stränge getrennt unter Verwendung von Denaturierungsbedingungen, wie beispielsweise einem hohen pH-Wert, Wärme oder Formamidlösungen. Die einzelnen DNA-Stränge werden sodann angewärmt in Duplexe. Während dieses Teils der DNA-Hybridisierung werden gemeinsame Subjektstränge in sr'-Hybridduplexe getrieben, und zwar mit dem gemeinsam verweilenden Überschuß komplementärer Referenzstränge. In Gegensatz dazu können die einzigartigen Glieder nicht in Hybriden eingefangen werden, da sie koresistente isogene Referenz-DNA's vermissen lassen. Statt dessen werden diese einzigartigen polymorphischen Stränge mit ihren Subjekt-Komplementen erwärmt. Überschüssige Referenzstränge bilden Duplexe mit ihren Komplementen. Die Erwärmungsreaktionsgeschwindigkeit vermindert sich, wenn die Bestandteile verbraucht werden und die Konzentration der Reaktanzien wird kleiner. Infolgedessen enthalten die Produkte des Erwärmungsschrittes Hybrid-DNA-Duplexe, reine Subjekt-DNA-Duplexe, reine Referenz-DNA-Duplexe und restliche Einzelsubjekt- und Referenz-DNA-Stränge. Partielle oder vollständige Zusammenfassung (poolings) der Nach-Hybridisierungsfractionen können ausgeführt werden in Vorbereitung für den Schritt sechs des Verfahrens, da das Poolen der DNA in Größenklassen die spätere Recombinations-DNA-Formation(en) erleichtert.

Im bevorzugten Ausführungsbeispiel erfolgt die DNA-Hybridisierung vorzugsweise in situ, d. h. in dem während der Fraktionierung verwendeten Agarosegel. Die in situ Hybridisierung bewahrt in maximaler Weise die in Schritt vier des Prozesses erreichten Fraktionierungen und begünstigt die Bildung von DNA-Duplexen

ständigen Reinigung der PRF-Glieder, welche den Polymorphismus repräsentieren.

Weitere Vorteile, Ziele und Einzelheiten der Erfindung ergeben sich aus der Beschreibung von Ausführungsbeispielen anhand der Zeichnung. In der Fig. 1 ist ein Prozeßflußdiagramm dargestellt, welches ein Verfahren zur partiellen Reinigung einzigartiger PRF-Konstituenten oder Bestandteile zeigt.

Im folgenden seien bevorzugte Ausführungsbeispiele der Erfindung beschrieben.

Die Zeichnung zeigt ein Prozeßflußdiagramm und illustriert eine Form der Erfindung für den subtraktiven Vergleich von zwei PRF's, welche die genomische DNA von zwei in Beziehung stehenden (verwandten) Genomen repräsentieren und für die partielle Reinigung der Restriktionsfragmente der subjektgenomischen DNA, die in der PRF der referenzgenomischen DNA nicht repräsentiert sind. Beim Anfangsschritt des Verfahrens wird eines der zwei in Beziehung stehenden Genome als das Subjektgenom und das andere als das Referenzgenom bezeichnet. Die genomische DNA der zwei Genome wird in ihrer entsprechenden Subjekt- und Referenzpopulation repräsentativer Fragmente (im folgenden als Subjekt- und Referenz-PRF's bezeichnet) umgewandelt. Die Umwandlung jeder genomischen DNA in ein PRF wird erreicht durch Trennung der Duplex-DNA-Moleküle mit irgendeiner konventionellen Restriktionsnuklease oder durch irgendein anderes Trenn- oder Schneidverfahren, welches definierte Fragmente ergibt im Gegensatz zu zufällig gebrochenen oder zerteilten Strängen.

Im bevorzugten Ausführungsbeispiel wird die subjekt- und referenzgenomische DNA im Anfangsschritt in ihrer entsprechenden PRF's umgewandelt unter Verwendung konventioneller isoschizomerischer Nukleasen. Restriktionsnukleasen, welche die gleiche Target- oder Zielsequenz in einem doppel-strängigen DNA-Molekül erkennen, aber die Stränge an unterschiedlichen Basenresten schneiden oder trennen, sind Isoschizomere voneinander. Beispielsweise erkennt die konventionelle Restriktionsnuklease Asp718 die Sechs-Basenpaar-Targetsequenz in Duplex DNA-Molekülen:

5' --- GGTACC --- 3'  
3' --- CCATGG --- 5'

wobei die gestrichelte Linie Nicht-Targetteile der DNA-Stränge repräsentiert. Die Asp718 schneidet die Stränge des DNA-Moleküls zwischen zwei G's. Das Isoschizomer von Asp718, nämlich KpnI, erkennt die gleichen Basissequenzen, aber schneidet identische Stränge des DNA-Moleküls zwischen zwei C's. Das Schneiden einer von zwei identischen Proben von genomischer DNA mit Asp718 und der anderen mit KpnI erzeugt PRF's, die identische Mitgliedschaft haben, aber mit unterschiedlichen Fragmentenden. Die mit Asp718 erzeugten PRF's

5' --- G                      3'  
3' --- CCATG                5'

Fragmentenden, wohingegen die mit KpnI erzeugten PRF's

5' --- GGTAC                3'  
3' --- C                      5'

Fragmentenden besitzen.

Im bevorzugten Ausführungsbeispiel für DNA wird die Subjekt-PRF mit Asp718 erzeugt und die Referenz-PRF wird mit KpnI (oder umgekehrt) erzeugt. Der Vorteil der Verwendung von Isoschizomeren besteht darin, daß sie die PRF's markieren, und zwar durch Vorsehen von Fragmentendifferenzen, welche die Identifikation einer Fragmentart bei Vorhandensein von anderen ermöglichen. Die Verwendung von Isoschizomeren ermöglicht auch die selektive Recombinant DNA Cloning (selektives Kloning mit rekombinierter DNA [oder DNS]) am Ende des Prozesses. Das spezielle Paar von Isoschizomeren, Asp718 und KpnI, wird aus zwei Gründen gewählt. Als erstes werden die charakteristischen Fragmentenden ohne weiteres als diskriminierende Endmarkierungen in späteren Schritten des erfindungsgemäßen Verfahrens verwendet. Zweitens ergibt die Sechs-Basenpaar-Targetgröße eine PRF-Mitgliedschaft mit brauchbarer Größenverteilung für Primärlängen-Fraktionierungen. Es ist zweckmäßig, die Duplex-Subjektfragmente als "ss" und die Duplex-Referenzfragmente als "rr" zu bezeichnen.

Der zweite Schritt des in Fig. 1 gezeigten Verfahrens besteht in der Herstellung oder Vorbereitung der Referenz-PRF des Schritts eins mit weiteren biochemischen und/oder isotopischen Markierungen. Referenz-Fragmente mit den zusätzlichen Markierungen werden als r'r' bezeichnet. Die Wahl dieser Markierungen wird eingeschränkt durch das Erfordernis, daß genetisch identische ss-, rr- und r'r'-Fragmente identische Beweglichkeiten während des darauffolgenden primären Fraktionierungsschritts, des Schritts Nr. 4 haben müssen. Die im zweiten Schritt hinzugegebenen Markierungen ermöglichen die Trennung der Referenz-DNA's und DNA-Hybride von Subjekt- und Referenz-DNA-Strängen von Subjekt-DNA's während der sekundären Fraktionierung im Schritt sechs.

Im bevorzugten Ausführungsbeispiel umfaßt der zweite Schritt die photodynamische Biotinylation der Referenz-Restriktions-Fragmente, d. h. die Verwendung eines Lichtverfahrens zur Hinzufügung einer Biotinmarkierung zu den Referenz-Fragmenten. Dieser Schritt ergibt eine zweite Art der Markierung des Referenz-PRF's. Biotinylierte DNA's können stark mit einem Chromatographieharz gebunden werden, mit angebrachtem Avidin oder Streptavidin. Solche Chromatographieschritte ermöglichen die Zurückhaltung der biotinylierten